Journal of Mathematical Analysis and Applications **249,** 252–260 (2000)
doi:10.1006/jmaa.2000.6932, available online at http://www.idealibrary.com on IDEAL®

A Divisibility Theorem For Factorials

Harold N. Shapiro

Courant Institute of Mathematical Sciences, 251 Mercer Street

and

Gerson H. Sparer

Pratt Institute, Brooklyn, New York 11205

Received January 4, 2000

1. INTRODUCTION

Let $P(n)$ denote the largest prime factor of n, and set $\mathscr{B}(x)$ = the set of integers $n \le x$ such that n does not divide $P(n)!$ (i.e., $P(n)$ factorial). In [1] Erdos proposed that $|\mathscr{B}(x)| = o(x)$. A solution was given in [2]. This was improved to $|\mathscr{B}(x)| = O(x \exp(-c\sqrt{\log x}))$ for some constant $c > 0$ [3]. In a sense this says something about the relative manner in which prime powers divide the integers. We propose to refine this in a different direction.

A first thought is to consider the possibility that a power of n, n^k, $k > 1$, divides $P(n)!$. But this fails since $P(n)$ divides $P(n)!$ exactly to the first power. This suggests omitting the largest prime from the factorization of n, and introducing

$$\tilde{n} = \prod_{p \mid n,\, p < P(n)} p^{\nu_p(n)}, \tag{1.1}$$

where $\nu_p(n)$ = the exponent of the highest power of p which divides n. For any function $f(x) \ge 1$ we consider

$$\mathscr{S}(f, x) = \{n \le x, \tilde{n}^{f(x)} \text{ does not divide } P(n)!\}, \tag{1.2}$$

0022-247X/00 $35.00

and wish to characterize those $f(x)$ such that

$$|\mathscr{S}(f, x)| = o(x). \tag{1.3}$$

($|\mathscr{S}|$ denotes the number of elements in the set $\mathscr{S}$.)

DEFINITION 1.1. The function $f(x)$ belongs to the lower class $\mathscr{L}$ if (1.3) holds. Otherwise $f(x) \in \mathscr{U}$, the upper class.

Remark. $f(x) \in \mathscr{U}$ is equivalent to the existence of a constant $c > 0$ and a sequence of $x_i \to \infty$ such that $|\mathscr{S}(f, x_i)| \geq c\, x_i$ for all i.

We pass to a logarithmic scale and introduce

$$\varepsilon(f, x) = \varepsilon(x) = \frac{\log f(x)}{\log x}. \tag{1.4}$$

Our main objective is to establish

THEOREM 1.1. *A necessary and sufficient condition for $f(x) \in \mathscr{L}$ is that*

$$\lim_{x \to \infty} \varepsilon(f, x) = 0. \tag{1.5}$$

(*Hence when* (1.5) *fails* $f(x) \in \mathscr{U}$.)

2. THE NECESSITY OF (1.5)

Assuming that (1.5) fails, we wish to show that $f(x) \in \mathscr{U}$. Then

$$\overline{\lim_{x \to \infty}}\, \varepsilon(f, x) > 0, \tag{2.1}$$

so that on some sequence $x_i \to \infty$, $\varepsilon(f, x_i) > c > 0$. Thus it suffices to show that there exists a constant $d > 0$ such that

$$|\mathscr{S}(f, x_i)| > d x_i. \tag{2.2}$$

An integer $n \in \mathscr{S}(f, x)$ if

$$f(x)\nu_p(n) > \nu_p(P(n)!) \tag{2.3}$$

for some prime $p \mid n$, $p < P(n)$. Since $\nu_p(n) \geq 1$ and $\nu_p(P(n)!) \leq P(n)$, (2.3) is implied by

$$f(x) \geq P(n). \tag{2.4}$$

The set of $n \leq x$ satisfying (2.4) is contained in $\mathscr{S}(f, x)$ and it suffices to show that for $x = x_i$, (2.4) is satisfied by more than $d\, x_i$ values of $n \leq x_i$.

This is an immediate consequence of the following

LEMMA 2.1. *For any fixed constant $c > 0$, there exists a constant $d > 0$, depending on c, such that the number of $n \leq x$ for which*

$$P(n) \leq x^c \tag{2.5}$$

is greater than $d\,x$, for all sufficiently large x.

(The following proof was suggested by P. Erdos. In a later section we'll consider the actual density, $\delta(c)$, of these integers, which provides another proof as well as a method for calculating $\delta(c)$.)

Proof. It clearly suffices to consider the cases where $c = \frac{1}{k}$, k an integer ≥ 2. Choose ε, $0 < \varepsilon < \frac{1}{k+1}$, and consider the integers $n \leq x$ which are divisible by an integer of the form

$$m = p_1 p_2 \cdots p_k, \qquad x^{(1-\varepsilon)\frac{1}{k}} < p_i \leq x^{\frac{1}{k}}, i = 1, \dots, k, \tag{2.6}$$

where the p_i are primes. Such an integer is of the form tm where $1 \leq t < x^{\varepsilon}$, so that all the prime factors of t are less than $x^{\varepsilon} < x^{1/k}$. Further, our choice of ε implies that $\varepsilon < (1 - \varepsilon)\frac{1}{k}$, so that the representations in the form tm are unique. Thus, the integers divisible by those of the form (2.6) satisfy the condition (2.5) and have a count of at least

$$\begin{aligned}\frac{1}{k!}\,\frac{x}{2} \sum_{x^{\frac{1-\varepsilon}{k}} < p_i \leq x^{\frac{1}{k}}} \frac{1}{p_1 \cdots p_k} &= \frac{x}{2k!}\left(\sum_{x^{\frac{1-\varepsilon}{k}} < p \leq x^{\frac{1}{k}}} \frac{1}{p}\right)^k \\ &= \frac{x}{2k!}\left(\log \frac{1}{1-\varepsilon} + o(1)\right)^k\end{aligned}$$

and the lemma follows.

3. THE SUFFICIENCY OF (1.5)

Our next objective is to show that (1.5) implies that the number of $n \leq x$ such that (2.3) holds, for some prime $p < P(n)$ which divides n, is $o(x)$. Here we must replace $\mathscr{S}(f, x)$ by an appropriate larger set. Also, without loss of generality we may assume that $f(x) \to \infty$ as $x \to \infty$. Since $\nu_p(P(n)!) \geq [\frac{P(n)}{p}] > \frac{P(n)}{2p}$, (2.3) implies

$$\nu_p(n) f(x) > \frac{P(n)}{2p}, \tag{3.1}$$

for some prime p which divides n.

The set of $n \le x$ such that $\nu_p(n) \ge T$, for some prime p, is at most

$$x \sum_p \left\{ \frac{1}{p^T} + \frac{1}{p^{T+1}} + \cdots \right\} = x \sum_p \frac{1}{p^T} \left(1 - \frac{1}{p} \right)^{-1}$$
$$\le 2x \sum_{n=2}^{\infty} \frac{1}{n^T} = O\left(x \int_2^{\infty} \frac{du}{u^T} \right) = O\left(\frac{x}{T} \right),$$

which is $o(x)$ for $T \to \infty$ however slowly. Thus the set of interest is contained in the $n \le x$ such that

$$Tf(x) > \frac{P(n)}{2p}, \tag{3.2}$$

for some prime $p < P(n)$ which divides n. But such an n is divisible by a prime p such that

$$P(n) > p > \frac{P(n)}{2Tf(x)}. \tag{3.3}$$

We propose to show that the number of such n is $o(x)$.

Since $P(n)$ is the largest prime dividing n, these n are relatively prime to all primes less than x and greater than $P(n)$. We then have two cases:

Case I. $P(n) < x^{\varepsilon_1(x)}$ where $\varepsilon_1(x) \to 0$ is chosen later.

Then the number of $n \le x$ which are prime to all primes $p > P(n)$ is at most the number of $n \le x$ which are prime to all $p > x^{\varepsilon_1(x)}$. Next choose an $\varepsilon_2(x) > \varepsilon_1(x)$, which also tends to 0, but more slowly than ε_1, that is, so that $\varepsilon_1(x)/\varepsilon_2(x) \to 0$ as $x \to \infty$. Hence the integers under consideration are prime to all p in the interval $x^{\varepsilon_1(x)} < p < x^{\varepsilon_2(x)}$. It is a well known sieving result [4] that this number is

$$O\left(x \prod_{x^{\varepsilon_1} < p \le x^{\varepsilon_2}} \left(1 - \frac{1}{p} \right) \right) = O\left(x \frac{\varepsilon_1(x)}{\varepsilon_2(x)} \right),$$

which is $o(x)$.

Case II. $P(n) > x^{\varepsilon_1(x)}$, where $\varepsilon_1(x) \to 0$ (as adjusted again later).

We choose $\varepsilon_1(x) \to 0$ so that $\varepsilon(x)/\varepsilon_1(x) \to 0$ and $x^{\varepsilon_1}/f(x) = x^{\varepsilon_1 - \varepsilon} \to \infty$. Note that this implies that $x^{\varepsilon_1}/Tf(x) \to \infty$ if $T \to \infty$ sufficiently slowly.

Then the $n \le x$ under consideration are divisible by a prime $q = P(n)$, and by another prime p, where

$$\frac{q}{2Tf(x)} < p < q. \tag{3.4}$$

The number of such n is at most

$$x \sum_{x^{\varepsilon_1} < q \le x} \frac{1}{q} \sum_{\frac{q}{2Tf(x)} < p < q} \frac{1}{p}. \tag{3.5}$$

Using the fact that

$$\sum_{p \le z} \frac{1}{p} = \log\log z + C + O\left(\frac{1}{\log z}\right)$$

(for some constant C), we have

$$\sum_{\frac{q}{2Tf(x)} < p < q} \frac{1}{p} = \log\log q - \log\log \frac{q}{2Tf(x)} + O\left(\left(\log \frac{q}{2Tf(x)}\right)^{-1}\right). \tag{3.6}$$

The first two terms on the right of (3.6) equal

$$\log \frac{\log q}{\log\left(q/2Tf(x)\right)} = O\left(\frac{\log 2Tf(x)}{\log q}\right),$$

and the O term

$$\frac{1}{\log\left(q/2Tf(x)\right)} = O\left(\frac{\log 2Tf(x)}{\log q}\right).$$

Hence (3.5) equals

$$O\left(x \sum_{x^{\varepsilon_1} < q \le x} \frac{\log 2Tf(x)}{q \log q}\right) = O\left(x \frac{\log 2Tf(x)}{\varepsilon_1 \log x}\right).$$

Also choosing T so that $\log T = O(\varepsilon \log x)$, the above equals $O(x(\varepsilon/\varepsilon_1)) = o(x)$, since $\varepsilon/\varepsilon_1 \to 0$ as $x \to \infty$. (Note that this uses the fact that $f(x) \to \infty$.)

This completes the proof of Theorem 1.1.

4. ON THE VALUE OF $\delta(c)$

We return now to a detailed look at the value of the density $\delta(c)$ (whose existence has not as yet been demonstrated here). Recall that this is the density of integers such that (2.5) holds. That is, those integers $n \le x$, all of whose prime factors are less than or equal to x^c.

Clearly we may assume that $c < 1$. Also, such n are precisely those which are relatively prime to all primes p such that

$$x^c < p \le x. \tag{4.1}$$

Choose the integer $k \ge 1$ where

$$\frac{1}{k+1} \le c < \frac{1}{k}. \tag{4.2}$$

The density of the complement of the set of integers described above (if it exists),

$$\bar{\delta}(c) = 1 - \delta(c) \tag{4.3}$$

is given by

$$\bar{\delta}_k(c) = \bar{\delta}(c) = \sum_{i=1}^{k} (-1)^{i+1} \lim_{x\to\infty} \sum_{\substack{x^c < p_{j_1} < \cdots < p_{j_i} \le x \\ \prod_{\lambda=1}^{i} p_{j_\lambda} \le x}} (p_{j_1} \cdots p_{j_i})^{-1}. \tag{4.4}$$

Setting $G_i(c)$ equal to the last summation in (4.4) and $F_i(c)$ equal to the limit of this sum, we have

$$\bar{\delta}(c) = \sum_{i=1}^{k} (-1)^{i+1} F_i(c). \tag{4.5}$$

LEMMA 4.1. *For any fixed c*, $0 < c < 1$ (*recall* (4.2)), *for* $1 \le m \le k$,

$$F_m(c) = \int \underset{\mathscr{D}_m}{\cdots} \int \prod_{i=1}^{m} \frac{dx_i}{x_i}, \tag{4.6}$$

where the domain $\mathscr{D}_m = \mathscr{D}_m(c)$ *is given by*

$$\mathscr{D}_m = \left\{ c < x_1 < x_2 < \cdots < x_m, \sum_{i=1}^{m} x_i \le 1 \right\}. \tag{4.7}$$

(*Note that for* $m > k$, $mc > 1$, *and* $F_m(c) = 0$.)

Proof. By the Prime Number Theorem [4], if p_i is the ith prime $p_i = (i \log i)(1 + o(1))$, then

$$G_m(c) = \sum_{\mathscr{R}_m} \left[\prod_{j=1}^{m} (i_j \log i_j)^{-1} (1 + o(1)) \right], \tag{4.8}$$

where the domain $\mathscr{R}_m$ is given by

$$\mathscr{R}_m = \left(\begin{array}{c} c < \left(\dfrac{\log i_1}{\log x}\right)(1 + o(1)) < \cdots < \left(\dfrac{\log i_m}{\log m}\right)(1 + o(1)) \le 1 \\ \Sigma_{j=1}^{m} \left(\dfrac{\log i_j}{\log x}\right)(1 + o(1)) \le 1 \end{array} \right). \tag{4.9}$$

Writing

$$\prod_{j=1}^{m} (i_j \log i_j)^{-1} = \prod_{j=1}^{m} \left(\frac{\log i_j}{\log x}\right)^{-1} (i_j \log x)^{-1}$$

we see that the right side of (4.9) is a Riemann sum for the integral in (4.6). (Note that the $(1 + o(1))$ factors are easily disposed of.)

LEMMA 4.2. *For c and k as specified in Lemma* 4.1, *and* $1 \le m \le k$,

$$F_m(c) = \int_c^{\frac{1}{m}} F_{m-1}\left(\frac{u}{1-u}\right) \frac{du}{u}. \tag{4.10}$$

(*Note that $F_0(c) \equiv 1$ is consistent with the above*.)

Proof. Clearly (4.10) holds for $m = 1$, and we assume $m \ge 2$. From (4.6)

$$F_m(c) = \int_c^{\frac{1}{m}} \frac{dx_1}{x_1} \int \underset{\mathscr{K}_{m-1}}{\cdots} \int \prod_{i=2}^{m} \frac{dx_i}{x_i}, \tag{4.11}$$

where $\mathscr{K}_{m-1}$ is the domain

$$\mathscr{K}_{m-1} = \mathscr{K}_{m-1}(x_1) = \left[x_1 < x_2 < \cdots < x_m, \sum_{i=2}^{m} x_i \le 1 - x_1 \right]. \tag{4.12}$$

Introducing new variables $y_i = x_i/(1 - x_1)$, $i = 2, \dots, m$, the domain of the inner integer becomes

$$\frac{x_1}{1 - x_1} < y_2 < \cdots < y_m, \sum_{i=2}^{m} y_i < 1,$$

and

$$\prod_{i=2}^{m} \frac{dx_i}{x_i} = \prod_{i=2}^{m} \frac{dy_i}{y_i}$$

so that (4.11) yields (4.10).

LEMMA 4.3. *For any* $c > 0$, $\frac{1}{k+1} \leq c < \frac{1}{k}$, *and all* $m = 1, \dots, k$, $F_m(c)$ *is differentiable at* c, *and*

$$c \frac{d}{dc} F_m(c) = -F_{m-1}\left(\frac{c}{1-c}\right). \tag{4.13}$$

Proof. Since $F_0(c) \equiv 1$ we have $F_1(c) = -\log c$, which has derivative $-\frac{1}{c}$ in the interval $\frac{1}{2} \leq c < 1$. Proceeding by induction we assume the lemma for $m - 1$ where $2 \leq m \leq k$. Since the mapping $u \to \frac{u}{1-u}$ takes the interval $[\frac{1}{k+1}, \frac{1}{k})$ onto $[\frac{1}{k}, \frac{1}{k-1})$, F_{m-1} is certainly continuous at $\frac{c}{1-c}$. Hence $F_m(c)$ has a derivative at c and (4.13) follows from (4.10).

LEMMA 4.4. *For* $0 < c < 1$, $\frac{1}{k+1} \leq c < \frac{1}{k}$, *we have*

$$-c \frac{d}{dc} \bar{\delta}_k(c) = 1 - \bar{\delta}_{k-1}\left(\frac{c}{1-c}\right). \tag{4.14}$$

Proof. Applying (4.5) and (4.13) for $m = 1, \dots, k$, we obtain

$$\begin{aligned} -c \frac{d}{dc} \bar{\delta}_k(c) &= \sum_{m=1}^{k} (-1)^m c \frac{d}{dc} F_m(c) \\ &= 1 - \sum_{m=2}^{k} (-1)^m F_{m-1}\left(\frac{c}{1-c}\right), \end{aligned}$$

and (4.14) follows.

COROLLARY. *The density* $\delta(c)$ *satisfies*

$$c \frac{d}{dc} \delta_k(c) = \delta_{k-1}\left(\frac{c}{1-c}\right). \tag{4.15}$$

Another Proof of Lemma 2.1. Consider first the case of $\frac{1}{2} < c \leq 1$. Then

$$\bar{\delta}(c) = \bar{\delta}_1(c) = F_1(c) = \log \frac{1}{c} < \log 2 < 1.$$

Assuming that we have a real c such that $\bar{\delta}_k(c) = 1$, it follows that $0 < c \leq \frac{1}{2}$. Setting

$$c^* = \limsup_{\delta(c)=1} c, \tag{4.16}$$

for all $c < c^*$ we have $\bar{\delta}(c) = 1$. Then if $\frac{1}{k+1} < c^* \le \frac{1}{k}$, we have $k \ge 2$, and there is an interval of c values $\frac{1}{k+1} < c < c^*$ such that $\bar{\delta}(c) = 1$ for all of these. Then $\frac{d}{dc}\bar{\delta}(c) = 0$, and (4.14) yields that $\bar{\delta}_{k-1}(\frac{c}{1-c}) = 1$. But $c^* \le \frac{1}{k} < \frac{c}{1-c}$, which contradicts (4.16).

Turning next to the calculation of $\delta(c)$, integrating (4.15) yields

$$\delta_k\left(\frac{1}{k}\right) - \delta_k(c) = \int_c^{\frac{1}{k}} \frac{1}{u}\delta_{k-1}\left(\frac{u}{1-u}\right) du. \tag{4.17}$$

Making a change of variable this becomes

$$\delta_k(c) = \delta_k\left(\frac{1}{k}\right) - \int_{\frac{c}{1-c}}^{\frac{1}{k-1}} \frac{1}{u(u+1)}\delta_{k-1}(u)\, du. \tag{4.18}$$

This last provides a natural numerical algorithm for calculating $\delta(c)$.

REFERENCES

1. P. Erdos, A proposed problem, *Amer. Math. Monthly* **98** (1991), 965.
2. I. Kastanas, The smallest factorial that is a multiple of n, *Amer. Math. Monthly* **101** (1994), 179.
3. S. Akbik, On a density theorem of Erdos, *J. Math. Math. Sci.* **22** (1999), 655–658.
4. H. Halberstam and K. F. Roth, "Sequences," Vol. I, Clarendon Press, Oxford, 1966.
5. A. O. Gelfond and Yu V. Linnik, "Elementary Methods in Analytic Number Theory" (translated by A. Feinstein), Rand McNally, Chicago, 1965.